

ABSTRAK

Stunting adalah masalah kesehatan kronis yang memengaruhi pertumbuhan anak dan balita. Dalam penelitian ini, wilayah Puskesmas Maja masih memiliki angka kasus *stunting* yang cukup tinggi. Penelitian ini bertujuan mendiagnosa *stunting* serta mengidentifikasi daerah rawan dengan menerapkan algoritma *Random Forest* yang dikombinasikan dengan teknik SMOTE serta variasinya seperti SMOTE-*Tomek Links* untuk menangani ketidakseimbangan data. Data yang digunakan berjumlah 16.515 dari tahun 2022–2024. Proses mencakup *preprocessing*, pelatihan model, dan evaluasi menggunakan metrik akurasi, presisi, *recall*, *F1-score*, dan AUC. Dari tiga skenario yang diuji, model *Random Forest* dengan SMOTE *Tomek Links* dipilih sebagai model terbaik karena menghasilkan performa paling seimbang, dengan akurasi 90,21%, *precision* 37,63%, *recall* 79,87%, *F1-score* 51,16%, dan AUC 93,12%. Model ini juga diterapkan dalam aplikasi GUI berbasis *Streamlit* untuk memudahkan prediksi *stunting* serta visualisasi daerah rawan *stunting*. Sistem ini diharapkan dapat mendukung tenaga kesehatan dalam deteksi dini *stunting* secara akurat dan efisien.

Kata Kunci: Diagnosa, *Random Forest*, SMOTE, SMOTE-*Tomek Links*, *Stunting*

ABSTRACT

Stunting is a chronic health issue that affects the growth and development of children and toddlers. In this study, the Maja Health Center area continues to report a relatively high prevalence of stunting cases. This research aims to diagnose stunting and identify high-risk areas by implementing the Random Forest algorithm, combined with the SMOTE technique and its variation, SMOTE-Tomek Links, to address data imbalance. A total of 16,515 data records from 2022 to 2024 were used. The process includes data preprocessing, model training, and evaluation using accuracy, precision, recall, F1-score, and AUC metrics. Among the three tested scenarios, the Random Forest model with SMOTE-Tomek Links was selected as the best-performing model, achieving balanced results with 90.21% accuracy, 37.63% precision, 79.87% recall, 51.16% F1-score, and 93.12% AUC. The model is also implemented in a Streamlit-based GUI application to facilitate stunting prediction and visualization of high-risk areas. This system is expected to assist healthcare workers in early and accurate stunting detection.

Keywords: Diagnosis, Random Forest, SMOTE, SMOTE-Tomek Links, Stunting